

王倩, 葛宏义, 蒋玉英, 等. 基于太赫兹时域光谱的大米品种识别研究 [J]. 食品工业科技, 2022, 43(23): 19–25. doi: 10.13386/j.issn1002-0306.2021120292

WANG Qian, GE Hongyi, JIANG Yuying, et al. Identification of Rice Varieties Based on Terahertz Time Domain Spectroscopy[J]. Science and Technology of Food Industry, 2022, 43(23): 19–25. (in Chinese with English abstract). doi: 10.13386/j.issn1002-0306.2021120292

· 未来食品 ·

基于太赫兹时域光谱的大米品种识别研究

王倩¹, 葛宏义^{1,2,*}, 蒋玉英^{1,2}, 张元^{1,2}, 秦一菲^{1,2}

(1.河南工业大学信息科学与工程学院, 河南郑州 450001;

2.河南工业大学粮食信息处理与控制教育部重点实验室, 河南郑州 450001)

摘要:为实现大米品种的准确鉴别, 提出一种基于太赫兹时域光谱 (Terahertz Time-Domain Spectroscopy, THz-TDS) 技术的大米品种识别方法。利用标准差 (Standard Deviation, SD) 和区间偏最小二乘 (Interval Partial Least Square, iPLS) 选取 0.53~1.21 THz 波段的吸收光谱信息作为分类模型的输入数据, 再采用决策树模型 (Decision Tree, DT) 对四种大米吸收光谱进行分类识别, 并在模型训练过程中结合网格搜索算法寻找模型最优参数。为增加实验对比度, 分别使用逻辑回归模型和支持向量机模型进行对比实验, 其模型分类准确率分别为 80.75% 和 88.75%。实验结果表明, 太赫兹时域光谱技术结合 SD、iPLS 和 DT 方法可以实现大米品种的准确识别, 准确率可达 95%, 为农产品品种识别提供了一种新的鉴别方法。

关键词: 太赫兹时域光谱, 大米品种, 标准差, 区间偏最小二乘, 决策树

中图分类号: TS211.7

文献标识码: A

文章编号: 1002-0306(2022)23-0019-07

DOI: 10.13386/j.issn1002-0306.2021120292



本文网刊:

Identification of Rice Varieties Based on Terahertz Time Domain Spectroscopy

WANG Qian¹, GE Hongyi^{1,2,*}, JIANG Yuying^{1,2}, ZHANG Yuan^{1,2}, QIN Yifei^{1,2}

(1.College of Information Science and Engineering, Henan University of Technology, Zhengzhou 450001, China;

2.Key Laboratory of Grain Information Processing and Control, Henan University of Technology, Ministry of Education, Zhengzhou 450001, China)

Abstract: In order to achieve accurate identification of rice varieties, a rice variety identification method based on terahertz time-domain spectroscopy (THz-TDS) technology was proposed. To select the terahertz band by Interval partial least square (iPLS) and standard deviation (SD), which determined the absorption spectrum data of the 0.53~1.21 THz band as the input data of the classification model. Then to use the decision tree (DT) identified absorption spectra of four kinds of rice and the model parameters was obtained by combining with the grid search algorithm. In order to increase the experimental contrast, logistic regression models and support vector machine models were used separately for comparative experiments, and the model classification accuracy was 80.75% and 88.75%, respectively. Experimental results showed that the terahertz time-domain spectroscopy technology combined with SD, iPLS and DT methods can realize the accurate identification of rice varieties with an accuracy rate of up to 95%, providing a new identification method for the identification of agricultural varieties.

Key words: terahertz time-domain spectroscopy(THz-TDS); rice varieties; standard deviation(SD); interval partial least squares(iPLS); decision tree(DT)

收稿日期: 2021-12-27

基金项目: 国家自然科学基金 (61975053)。

作者简介: 王倩 (1993-), 女, 硕士研究生, 研究方向: 太赫兹波检测, E-mail: wangqian20210411@163.com。

* 通信作者: 葛宏义 (1983-), 男, 博士, 副教授, 研究方向: THz 光谱与成像, E-mail: gehongyi2004@163.com。

随着人们生活品质的不断提高,消费者越来越热衷于购买地域性特色大米,如:黑龙江五常大米、河南原阳大米、宁夏珍珠糯米和云南遮放贡米等。不同产地的大米不仅具有特殊的品质和信誉,且经济价值也比普通大米高。不良商家为获取不当高额利润,通过对低价大米染色或工业加工等方式伪造特色大米,牟取暴利,造成大米市场混乱^[1],严重损害消费者和生产者权益。因此,亟需相应的检测方法识别大米品种,以规范大米市场。

针对大米品种识别方法较多,传统检测方法主要有感官评价法^[2]和理化指标技术法^[3]等,具有分析简单、方法直观等优势,但也存在一些不足,如感官评价法主观性强,对于采用化学剂染色大米或工业加工研磨大米,仅从外观上难以判别真伪;理化指标技术法操作繁琐、检测速度慢等。目前常见识别技术主要有电子舌技术^[4]、红外光谱技术^[5]和高光谱图像技术^[6],诸多学者以此为基础结合化学计量方法对大米识别做出研究,如惠延波等^[7]利用电子舌技术结合 PCA、DFA 两种模式对不同品种粉碎大米进行识别研究,区分指数 D1 为 95;王靖会等^[8]以高光谱成像技术为基础结合光谱和纹理特性实现不同大米品种的识别,识别率为 96.57%;王朝辉等^[9]利用近红外光谱技术结合 SG 平滑预处理方法和偏最小二乘判别分析方法对松原不同品种大米识别,识别率为 100%;刘亚超等^[10]利用近红外二维相关光谱对掺和和大米识别,识别率最高为 100%;虽然以上技术能有效解决传统技术检测精度不高、主观性强等缺点,但也存在一些不足,如电子舌技术检测结果易受传感器灵敏度和环境因素影响^[11]、设备集成度不高、价格昂贵、使用周期短等^[12];红外光谱技术灵敏度相对较低^[13]、检测结果易受各种环境影响^[14],红外光谱解释性差、化学计算依赖性高^[15];高光谱图像技术图像处理复杂性高、检测速度慢等^[16]。太赫兹(Terahertz, THz)^[17]是介于微波和红外之间的电磁波,频率在 0.1~10 THz 之间,具有穿透性强、电离辐射小、分子指纹等特点,还具有检测分析速度快、多成分同步分析、无损检测等优势,能有效弥补传统检测方法和目前常见技术的不足。其原理是利用飞秒脉冲产生 THz 电磁波,通过探测设备获得待测样品的光谱信息,再通过傅里叶变换获得测物品吸收和折射光谱信息。由于大分子振动和转动能级大多处于 THz 波段,并表现出明显吸收特性,因而可以通过光谱的指纹特性对大米品种进行识别。太赫兹时域光谱技术已被广泛应用到安全检查^[18]、生物医学^[19]、无线通信^[20]等领域,目前在农产品品质检测方面的应用也在不断增多,如农产品掺假^[21]、转基因作物检测^[22]和农产品主要成分含量检测^[23-24]等。太赫兹光谱技术的广泛应用,为大米品种识别提供了新思路。

本文采用太赫兹时域光谱技术结合标准差分析、区间偏最小二乘和决策树分类模型对不同大米

品种进行识别。分析了大米在光谱技术作用下的特点,建立了一种准确识别大米品种的分析模型,为大米品种识别提供了一种精准检测方法。

1 材料与方法

1.1 材料与仪器

实验样品分别是河南红米、珍珠糯米、黑米和富硒大米,均采购于永辉超市,如图 1 所示。大米颗粒利用研磨机研磨成粉末状,再利用电子称每次取 0.2 g 的粉末放入模具,并对模具使用 10 MPa 的压力施压 3 min,使粉末压制成型直径约为 13 mm、厚度约为 1.1 mm、表面均匀的圆形薄片。每种大米制备一组样品,每组样品制备 14 份,四组样品共 56 份。



图 1 实验样品

Fig.1 Experimental samples

太赫兹时域光谱仪 由河南工业大学教育部重点实验室提供,激光器波长 800 nm,光谱范围 0~3.5 THz,分辨率 0.03 THz,重复频率 80 MHz,脉冲波长 100 fs,信噪比 5000:1,具体系统结构如图 2 所示。

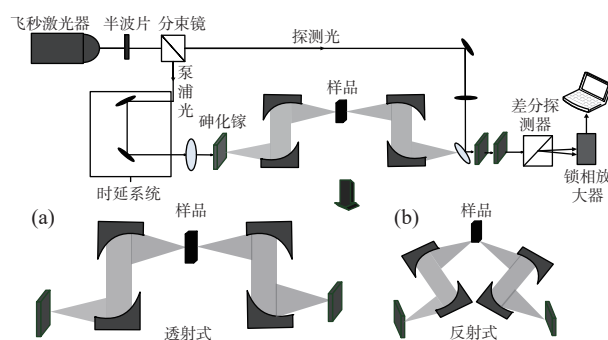


图 2 太赫兹时域光谱系统透射式(a)和反射式(b)

Fig.2 Terahertz time domain spectroscopy system transmission type (a) and reflection type (b)

时域光谱系统工作原理如下:飞秒激光器产生的光束作为输入光源,经分束镜分成能量较强的泵浦

光和能量相对弱的探测光。泵浦光先传输至时间延迟控制系统,再入射到光导天线或半导体晶体上激发出太赫兹脉冲,最后经两组抛物面镜聚焦到待测物体上。探测光经多次反射后和太赫兹脉冲共线触发太赫兹探测器,探测器通过对脉冲偏振状态的检测获得待测样品的时域波形。在整个测试过程中,由于水分对太赫兹脉冲有强吸收性,因此为减少空气中水分含量对实验结果的影响,需要不断对时域光谱系统补充氮气,以保持实验环境干燥。

1.2 实验步骤

实验分为三个阶段:第一阶段主要是对仪器参数和实验环境进行调整,光谱系统设置为透射模式,测试温度保持在 19 ℃ 左右,湿度保持在 2.8% 左右,并在测试正式开始前,仪器持续运行 2 min,以确保各项参数指标达到稳定状态;第二阶段是对样品进行检测,把制备好的样品放置在检测台旁,按标签顺序依次放入太赫兹时域光谱系统中进行检测,重复测量三次,并取平均值作为该样品时域光谱信息。为减少环境因素对实验结果的影响,取样间隔时间不超过 5 s;第三阶段是利用 Origin 和 Python 软件结合分类模型算法对获得的太赫兹光谱数据信息进行数据预处理和数据分类。

1.3 模型与建模方法

1.3.1 标准差 标准差(SD)值反映的是数据集的波动程度,即离散程度^[25]。标准差值大表示数据和平均值之间差异较大,标准差值小表示数据和平均值之间差异小,在光谱信息中标准差值大小表示光谱波动范围,且两者为成正比关系。标准差的计算公式如下:

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - u)^2} \quad \text{式 (1)}$$

式中: x_i 为样品数据; u 为算术平均值; N 为数值个数。

1.3.2 区间偏最小二乘 区间偏最小二乘(iPLS)主要用于光谱波段的选择,其特点是采用偏最小二乘实现光谱波段的局部回归分析,并根据局部回归分析的最小均方根误差确定光谱特征波段的选择。偏最小二乘的原理是将全光谱划分为多个具有同样宽度的区间,然后在各个子区间内使用偏最小二乘回归,根据各个子区间均方根误差的比较,选择均方根误差最小的子区间对应的光谱波段为光谱特征波段^[26]。

1.3.3 决策树 决策树(DT)模型是一种类似于二叉树的网络结构,具有可读性、分类速度快等优点,在分类、回归等方面有广泛的应用。决策树中每一个非叶子节点都是一个决策点,即判断条件,满足条件的放在节点的右侧,不满足的放在节点的左侧。如何确定决策点是构造决策树的重点,根据决策点特征选择的不同,常见的方法有 ID3 算法、C4.5 算法和 CART 算法。ID3 算法的核心是在决策树各个节点上应用信息增益准则选择特征,选择信息增益最大的

特征作为决策点。但 ID3 算法没有考虑到数据的连续性和过拟合问题, C4.5 算法在弥补 ID3 算法的不足的同时也提出了根据信息增益率作为特征选择的准则。C4.5 和 ID3 算法都能有效的解决分类问题,但不能处理回归问题,而根据 CART 算法构建的决策树不仅能处理分类问题也能做回归分析,在分类问题时选择基尼系数最小的特征作为决策点,在回归分析中 CART 算法使用平方误差最小值对应的特征作为决策点^[27]。

1.3.4 SD-iPLS-DT 由于原始数据集维度高、信噪比低,直接使用决策树模型进行分类处理,易造成数据处理速度慢,模型分类准确率低等问题。因此采用 SD-iPLS 与 DT 联用的方法对测试样品进行分类,即先采用 SD 和 iPLS 选择合适的太赫兹波段作为模型的输入数据,再利用 DT 模型进行分类识别。本次实验中原始数据维度为 220,通过 SD 选择稳定性较好的光谱波段,同时结合 iPLS 选择均方根误差最小的光谱区间,根据两者最优结果,选择其中的 38 维吸收光谱数据作为分类模型的输入数据,最后使用 DT 算法进行大米品种识别研究。

1.4 数据处理

由于光谱信息中包含实验样品的振幅和相位信息,因此可以将时域光谱信号利用快速傅里叶变换得到实验样品的频域信息,并结合 Dorney^[28] 和 Dubi-llaret^[29] 等提出的数据处理模型处理光谱数据,得到大米样品的折射率和吸收系数。计算公式如下:

$$n(\omega) = \frac{(\varphi_s(\omega) - \varphi_r(\omega))c}{\omega d} + 1 \quad \text{式 (2)}$$

式中: $n(\omega)$ 表示光谱折射率, n ; ω 代表角速度, rad/s; φ_s 、 φ_r 分别表示样品和参考信号的相位信息, p ; c 表示光谱传播速度, m/s; d 表示样品的厚度, mm。

$$\alpha(\omega) = \frac{2}{d} \ln \frac{4A_r n(\omega)}{A_s(n(\omega) + 1)^2} \quad \text{式 (3)}$$

式中: $\alpha(\omega)$ 表示光谱吸收系数, cm^{-1} ; A_r 、 A_s 分别表示参考信号和样品的振幅值, a.u。

2 结果与分析

2.1 光谱分析

2.1.1 时域和频域 为减少实验误差,对每份实验样品的时域信号分别测量 3 次,并取平均值作为样品的时域信号。图 3(a) 是样品的时域光谱,时域光谱信息经傅里叶变换得到样品的频谱信息,如图 3(b) 所示。可以看出样品信号相对于参考信号有一定的时延、衰减和重叠现象,产生时延可能是由于光谱穿透样用时较长,衰减可能是由样品表面的大颗粒物质对太赫兹光谱反射和内部样品的吸收造成的,而光谱之间区分度不明显可能是由于各种大米中的主要成分都是碳水化合物、蛋白质和脂肪。

2.1.2 折射率和吸收系数 太赫兹时域光谱数据中包含丰富的振幅和相位信息,经公式(2)和(3)计算得到样品的折射率和吸收系数,如图 4 所示。可以看

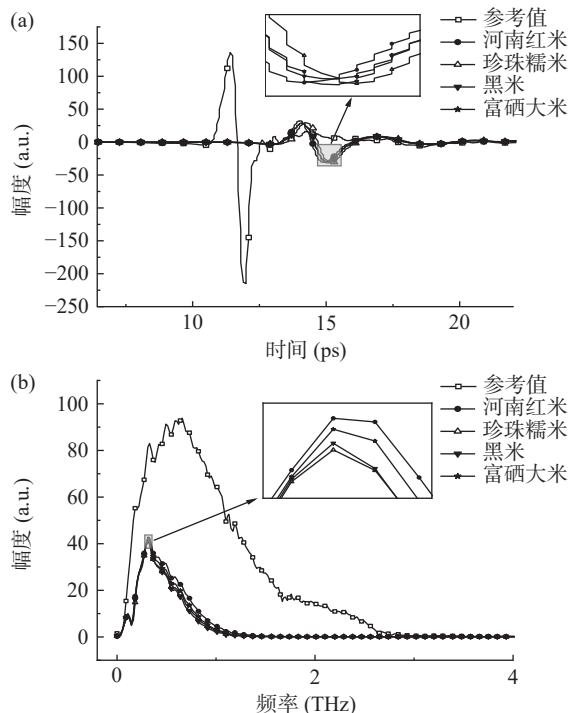


图3 时域光谱图(a)和频域光谱图(b)

Fig.3 Time domain spectroscopy (a) and frequency domain spectroscopy (b)

出四种样品的折射率曲线和吸收光谱曲线在低频波段均出现重叠现象,在高频波段折射率光谱曲线趋于一致,吸收光谱曲线区别明显,因此本文选择吸收光谱作为模型的输入数据。

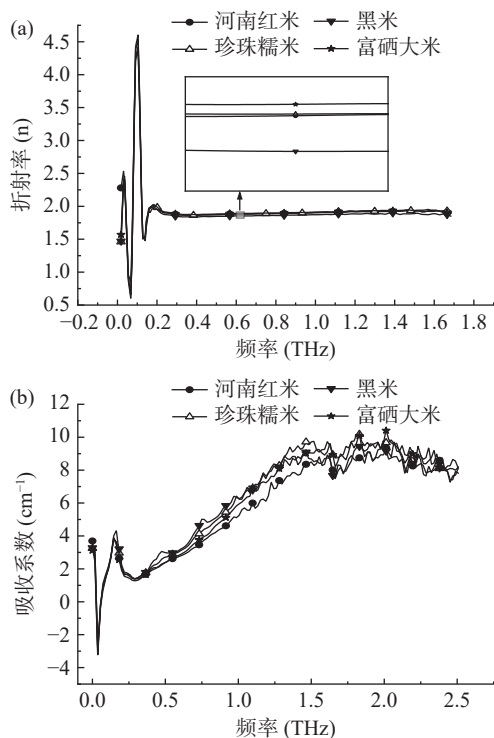


图4 折射光谱(a)和吸收光谱(b)

Fig.4 Refraction spectrum (a) and absorption spectrum (b)

2.2 光谱处理

由于吸收光谱存在信噪比低和光谱重叠现象,为更好识别大米品种,本文首先对吸收光谱数据进行

标准化处理解决光谱重叠问题,再利用标准差分析和区间偏最小二乘(iPIS)选取光谱稳定性好、信噪比高的太赫兹波段作为分类模型的输入数据。

2.2.1 吸收光谱预处理和稳定性分析 为增大吸收光谱曲线之间的差异性,本次实验采用标准化预处理方法对吸收光谱数据进行预处理操作。图5为预处理后四种样品的吸收光谱数据平均值,通过观察光谱曲线可以看出,四种样品的吸收光谱曲线差异明显。同时,利用光谱角度^[30]对光谱曲线之间的差异性进行评价,以河南红米吸收光谱为参照,结果如表1所示,可以看出在经过预处理之后,光谱角度的数值明显增大,即光谱曲线之间的差异性增大。但标准化预处理后的光谱仍存在信噪比低、光谱波动大等问题,因此为减少噪声和光谱波动性对模型准确率的影响,需选择合适的光谱波段作为模型分类识别的输入数据。

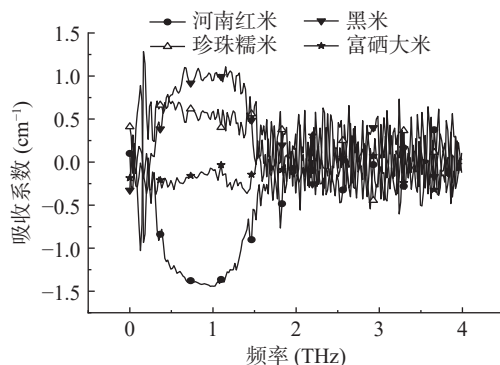


图5 吸收数据标准化预处理

Fig.5 Standardized preprocessing of absorption data

表1 不同样品光谱角度

Table 1 Different sample spectral angles

	红米与珍珠糯米	红米与黑米	红米与富硒大米
原始光谱角度	0.077	0.093	0.060
标准化预处理后	2.513	2.652	1.182

根据不同光谱波段的标准差对光谱的稳定性进行分析^[31],实验过程中把吸收全光谱划分为7个不同的区间,并对每个区间进行标准差分析,结果如表2、图6所示,可以看出0.53~1.21 THz之间四种大米样品标准差分别为0.06、0.07、0.08、0.06,平均

表2 四种样品预处理后不同波段标准差

Table 2 Standard deviation of different bands after pretreatment of four samples

太赫兹波段(THz)	河南红米样品	珍珠糯米样品	黑米样品	富硒大米样品	标准差平均值
0~0.53	0.65	0.53	0.36	0.12	0.42
0.53~1.21	0.06	0.07	0.08	0.06	0.07
1.21~1.74	0.40	0.19	0.45	0.18	0.31
1.74~2.31	0.29	0.19	0.25	0.20	0.23
2.31~2.71	0.28	0.28	0.27	0.25	0.27
2.71~3.40	0.25	0.30	0.30	0.23	0.27
3.40~4.00	0.24	0.25	0.28	0.20	0.24
0~4.00	0.59	0.38	0.45	0.21	0.41

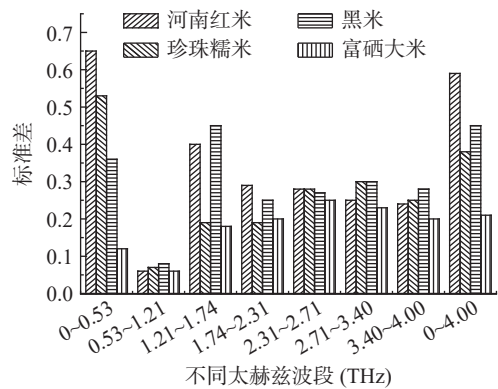


图 6 四种样品预处理后不同波段标准差

Fig.6 Standard deviation of different bands after pretreatment of four samples

值为 0.07, 是 8 组光谱数据中标准差值最小的一组, 即光谱波动最小, 稳定性最好的一组; 0~0.53 THz 之间的四种大米样品标准差分别为 0.65、0.53、0.36、0.12, 平均值为 0.42, 是 8 组光谱数据中标准差值最大的一组, 即光谱波动最大, 稳定性最差的一组。

2.2.2 吸收光谱特征谱区间选择 实验过程中采用 SPXY-iPLS 算法选择吸收光谱特征区间。首先利用 SPXY 算法^[32]进行数据集的划分, 再利用 iPLS 算法计算每个子区间的均方根误差。在子区间划分中, 把预处理后的全光谱分为 2、3、4、5、6、7、8、10、15 和 22 个子区间, 分别对各子区间建立 PLS 回归模型^[33], 进而比较模型的均方根误差, 确定均方根误差最小的回归模型对应的子区间。由表 3、图 7

表 3 iPLS 不同分割区间下最优区间列表

Table 3 List of optimal intervals under different segmentation intervals of interval partial least square

区间个数	最佳区间	最优因子	均方根误差	对应太赫兹波段(THz)
2	2	3	0.84	0~1.98
3	1	9	0.50	0~1.32
4	1	8	0.62	2.03~3.00
5	1	8	0.60	1.61~2.40
6	4	15	0.57	0.64~1.32
7	2	21	0.55	0.55~1.12
8	2	5	0.52	0.50~0.99
10	2	5	0.57	0.81~1.19
15	3	9	0.54	0.53~0.81
22	4	3	0.60	0.92~1.08

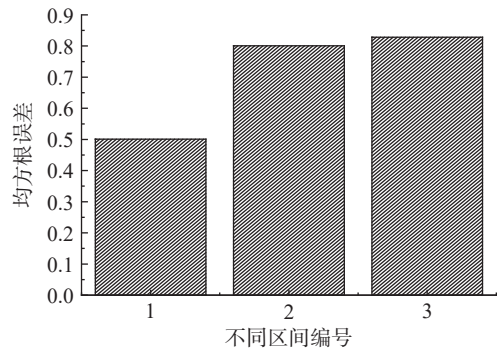


图 7 最优区间均方根误差

Fig.7 Optimal interval root mean square error

可以看出当划分区间个数为 2 时, 最佳区间对应的均方根 0.84, 是 10 组区间中均方根误差最大的区间; 当划分区间个数为 3 时, 最佳区间 1 区间对应的均方根误差为 0.50, 是 10 组区间中均方根误差最小的区间, 即模型效果最好, 对应的 THz 光谱波段为 0~1.32 THz。

结合光谱稳定性和区间偏最小二乘两者对应的最优波段, 最终选择 0.53~1.21 THz 之间的吸收光谱数据作为模型的输入数据。

2.3 模型分析

采用 SD-iPLS-DT 联用模型对样品进行分类。首先对数据集进行划分, 将 0.53~1.21 THz 之间的光谱数据按照 2:1 的比例划分为训练集和测试集。为实现模型最优分类结果, 结合网格搜索算法寻找最优参数组合。在决策树模型中主要的参数有特征选择评价标准(criterion)、最大深度(max_depth)。其中在决策树模型中特征评价标准主要参数有基尼系数(gini), 信息增益(entropy)。根据网格搜索算法最终确定模型的最优参数 criterion 为 gini、最大深度为 3。模型对应的分类结果如表 4 所示, 分类准确率为 95%。

表 4 决策树分类准确率

Table 4 Classification accuracy of decision tree

模型	测试集准确率(%)				
	河南红米	珍珠糯米	黑米	富硒大米	平均准确率
DT	100	80	100	100	95

为证明本文模型的性能, 将 SD-iPLS-DT 模型与常见的分类模型逻辑回归(LC)和支持向量机^[34](SVM)进行对比。其中在 SVM 模型中, 分别利用径向基核函数、线性核函数和 Sigmoid 核函数等三种不同核函数对样本进行分类。如表 5 所示, 可以看出在 SVM 模型中基于线性核函数的分类模型效果最好, 分类准确率可达 88.75%。由表 6 可以看出, 直接使用逻辑回归和支持向量机分类准确率分别为 80.75% 和 88.75%, 使用标准差和 iPLS 结合选择合适的光谱波段后, 再使用决策树模型进行分类准确率为 95%, 因此使用 SD-iPLS-DT 联用模型效果要优

表 5 不同核函数分类准确率

Table 5 Classification accuracy of different kernel functions

不同核函数	测试集准确率(%)				
	河南红米	珍珠糯米	黑米	富硒大米	平均准确率
RBF	100	60	85	25	67.5
linear	100	100	80	75	88.75
Sigmoid	100	40	80	25	61.25

表 6 不同模型的分类准确率

Table 6 Classification accuracy of different models

模型	平均准确率(%)
LC	80.75
SVM	88.75
SD-iPLS-DT	95

于逻辑回归和支持向量机模型。

3 结论与讨论

本文利用太赫兹时域光谱技术获取四种大米的时域和频域光谱数据,并计算得到折射率和吸收系数。采用标准差对吸收光谱数据稳定性进行分析,得出 0.53~1.21 THz 之间的光谱稳定性最好;区间偏最小二乘选择吸收光谱特征波段,确定 0~1.32 THz 区间的均方根误差最小。结合稳定性最好和最小均方根误差对应的太赫兹波段,选择 0.53~1.21 THz 波段吸收光谱信息作为决策树模型的输入数据。实验结果显示:所提出的 SD-iPLS-DT 的方法分类准确率可达 95%。为更好对比模型分类效果,利用逻辑回归、支持向量机与 SD-iPLS-DT 方法进行实验对比,实验结果显示:逻辑回归分类准确率为 80.75%,支持向量机准确率为 88.75%,SD-iPLS-DT 模型分类准确率为 95%。识别准确率高于现有利用近红外光谱技术进行稻花香大米定性分析研究 94% 准确率^[35]、利用高光谱成像技术对大米无损检测 92.96% 准确率^[36]、利用拉曼光谱大米产地识别 91.11% 准确率^[37]等。因此可以得出结论:时域光谱技术结合 SD-iPLS-DT 方法可以实现不同大米品种的准确识别。同时也为太赫兹时域光谱技术在农产品识别和质量安全检测方面提供了新的检测方法。

参考文献

- [1] 周广春,孟维初,全东兴,等.吉林省第八届优质食味水稻品种鉴评报告[J].东北农业科学,2018,43(6):1-4. [ZHOU G C, MENG W R, QUAN D X, et al. A report on the eighth session of evaluation of high eating quality rice varieties in Jilin Province[J]. Journal of Northeast Agricultural Sciences, 2018, 43(6): 1-4.]
- [2] 张玉荣,周显青,杨兰兰.大米食味品质评价方法的研究现状与展望[J].中国粮油学报,2009,24(8):155-160. [ZHANG Y R, ZHOU X Q, YANG L L. Present situation and expectation on methods for taste evaluation of rice[J]. Journal of the Chinese Cereals and Oils Association, 2009, 24(8): 155-160.]
- [3] CAMPOS I, ALCANIZ M, AGUADO D, et al. A voltammetric electronic tongue as tool for water quality monitoring in wastewater treatment plants[J]. Water Research, 2012, 46(8): 2605-2614.
- [4] 刘敏,谭书明,张洪礼,刘芳宏.不同品种大米口感品质分析[J].食品科学,2018,39(15):88-92. [LIU M, TAN S M, ZHANG H L, et al. Taste quality of different rice varieties[J]. Food Science, 2018, 39(15): 88-92.]
- [5] 刘晓欢,刘翠玲,孙晓荣,等.基于傅里叶红外光谱技术的大米产地溯源快速判别方法研究[J].食品科技,2021,46(4):244-249. [LIU X H, LIU C L, SUN X R, et al. Study on fast identification method of rice origin traceability based on fourier transform infrared spectroscopy technology[J]. Food Science and Technology, 2021, 46(4): 244-249.]
- [6] 赵凡,闫昭如,薛建新,等.高光谱无损识别野生和种植黑枸杞[J].光谱学与光谱分析,2021,41(1):201-205. [ZHAO F, YAN Z R, XUE J X, et al. Identification of wild black and cultivated Goji berries by hyperspectral Image[J]. Spectroscopy and Spectral Analysis, 2021, 41(1): 201-205.]
- [7] 惠延波,潘印卿,王莉,等.基于电子舌技术的不同品种大米区分识别研究[J].科学技术与工程,2014,14(33):250-254,259. [HUI Y B, PAN Y Q, WANG L, et al. Discrimination of different varieties of rice base on electronic tongue technique[J]. Science Technology and Engineering, 2014, 14(33): 250-254, 259.]
- [8] 王靖会,程娇娇,刘洋,等.基于高光谱成像技术鉴别大米品种[J].中国农业科技导报,2021,23(9):121-128. [WANG J H, CHEN J J, LIU Y, et al. Identification of rice variety based on hyperspectral imaging technology[J]. Journal of Agricultural Science and Technology, 2021, 23(9): 121-128.]
- [9] 王朝辉,高地,赖瀚清,等.近红外光谱法对松原不同品种大米的确定探究[J].食品研究与开发,2020,41(4):146-155. [WANG C H, GAO D, LAI H Q, et al. Confirmation of different rice varieties in Songyuan by near infrared spectroscopy[J]. Food Research and Development, 2020, 41(4): 146-155.]
- [10] 刘亚超,李永玉,彭彦昆,等.近红外二维相关光谱的掺和大米判别[J].光谱学与光谱分析,2020,40(5):1559-1564. [LIU Y C, LI Y Y, PENG Y K, et al. Application of two-dimensional correlation spectra in the identification of adulterated rice[J]. Spectroscopy and Spectral Analysis, 2020, 40(5): 1559-1564.]
- [11] 黄星奕,张浩玉,赵杰文.电子舌技术在食品领域应用研究进展[J].食品科技,2007(7):20-24. [HUANG X Y, ZHANG H Y, ZHAO J W. Research and application of electronic tongue technology in food industry[J]. Food Science and Technology, 2007(7): 20-24.]
- [12] 云雯,吴彦蕾,杨彦,等.电子舌在食品安全检测领域的研究进展[J].中国调味品,2014,39(7):133-137. [YUN W, WU Y L, YANG Y, et al. Research progress of application of electronic tongue technique in food safety testing[J]. China Condiment, 2014, 39(7): 133-137.]
- [13] 曾凤仙,周道志.红外光谱技术及其在食品检测中的应用[J].现代食品,2018(16):131-133. [ZENG F X, ZHOU D Z. Infrared spectroscopy and its application in food detection[J]. Modern Food, 2018(16): 131-133.]
- [14] 陈超,柳琦,李钊,等.红外光谱技术在食品安全检测中的应用与研究[J].食品研究与开发,2019,40(14):219-224. [CHEN C, LIU Q, LI F, et al. Research and application of infrared spectroscopy technology in food safety testing[J]. Food Research and Development, 2019, 40(14): 219-224.]
- [15] 曾子琦,蒋立文,刘霞,等.近红外光谱无损检测在食用油脂分析中的应用研究进展[J].中国油脂,2018,43(8):137-142. [ZENG Z Q, JING L W, LIU X, et al. Advance in application of near infrared spectroscopy in nondestructive analysis of edible oil[J]. China Oils and Fats, 2018, 43(8): 137-142.]
- [16] 滕安国,高峰,夏新成,等.高光谱技术在农业中的应用研究进展[J].江苏农业科学,2009(3):8-11. [TENG A G, GAO F, XIA X C, et al. Research progress on the application of hyperspectral technology in agriculture[J]. Jiangsu Agricultural Sciences, 2009(3): 8-11.]
- [17] MANJAPPA M, SINGH R. Materials for terahertz optical science and technology[J]. Advanced Optical Materials, 2020, 8(3): 1901984.
- [18] LI T, ZHANG L, HE J A, et al. Terahertz time-domain spec-

- troscopy for identification of hazardous substances in mail[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(12): 3724–3730.
- [19] LIU X H, ZHAO H W, LIU G F, et al. Application of terahertz technology in pharmaceutical setting[J]. *Progress in Chemistry*, 2010, 22(11): 2191–2198.
- [20] OUYANG A G, ZHANG Y L, LI B, et al. Study on the detection method of alum content in sweet potato starch by terahertz spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(3): 727–732.
- [21] 刘燕德, 杜秀洋, 李斌, 等. 太赫兹时域光谱技术对紫米掺假的检测研究[J]. *光谱学与光谱分析*, 2020, 40(8): 2382–2387. [LIU Y D, DU X Y, LI B, et al. Detection of purple rice adulteration by terahertz time domain spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(8): 2382–2387.]
- [22] JU X G, ZHANG Y, LIAN F Y, et al. Quick test for transgenic components in rice using terahertz spectra[J]. *Applied Spectroscopy*, 2019, 73(2): 171–181.
- [23] 孙彤, 张卓勇, 相玉红, 等. 利用 Far-IR 和 THz-TDS 光谱法测定大米中西维因[J]. *光谱学与光谱分析*, 2016, 36(2): 541–544. [SUN T, ZHANG Z Y, XIANG Y H, et al. Determination of carbaryl in rice by using Far-IR and THz-TDS techniques[J]. *Spectroscopy and Spectral Analysis*, 2016, 36(2): 541–544.]
- [24] JIANG Y Y, GE H Y, ZHANG Y. Quantitative determination of maltose concentration in wheat by using terahertz imaging[J]. *Spectroscopy and Spectral Analysis*, 2018, 38(10): 3017–3022.
- [25] 耿响, 张泽栋, 江龙发, 等. 基于近红外光谱技术的油茶籽粕中灰分含量快速检测方法[J]. *红外*, 2022, 43(1): 43–48. [GENG X, ZHANG Z D, JIANG L F, et al. A fast detection method of ash content in camellia seed meal based on near infrared spectroscopy technology[J]. *Infrared*, 2022, 43(1): 43–48.]
- [26] 张优优, 陈伟豪, 唐志敏, 等. 区间偏最小二乘结合差分进化算法应用于鱼粉近红外光谱波长筛选[J]. *分析测试学报*, 2020, 39(11): 1392–1397. [ZHANG Y Y, CHEN W H, TANG Z M, et al. Application of interval partial least squares with differential evolution algorithm in wavelength selection of near infrared spectroscopy for fishmeal[J]. *Journal of Instrumental Analysis*, 2020, 39(11): 1392–1397.]
- [27] 张坤. 决策树 CART 在高校学生体质测试数据分析中的应用研究[J]. *四川体育科学*, 2022, 41(1): 41–45. [ZHANG K. Application of decision tree CART in data analysis of physical fitness test of college students[J]. *Sichuan Sports Science*, 2022, 41(1): 41–45.]
- [28] DORNEY T D, BARANIUK R G, MITTLEMAN D M. Material parameter estimation with terahertz time-domain spectroscopy[J]. *Journal of the Optical Society of America*, 2001, 18(7): 1562–1571.
- [29] DUVILLARET L, GARET F, COUTAZ J L. Highly precise determination of optical constants and sample thickness in terahertz time-domain spectroscopy[J]. *Applied Optics*, 1999, 38(2): 409–415.
- [30] 张昊丹, 孙孝林, 王晓晴, 等. 可见-近红外光谱原位估测土壤有机质含量中测试点和采样位置导致的误差分析[J]. *光谱学与光谱分析*, 2020, 40(11): 3499–3507. [ZHANG H D, SUN X L, WANG X Q, et al. Analyzing errors due to measurement positions and sampling locations for in situ measurements of soil organic matter using vis-nir spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(11): 3499–3507.]
- [31] 熊岑, 张若秋, 李辉, 等. 最小角回归算法 (LAR) 结合采样误差分布分析 (SEPA) 建立稳健的近红外光谱分析模型[J]. *分析测试学报*, 2018, 37(7): 778–783. [XIONG Q, ZHANG R Q, LI H, et al. A robust near infrared modeling by least angel regression and sampling error profile analysis[J]. *Journal of Instrumental Analysis*, 2018, 37(7): 778–783.]
- [32] 王世芳, 韩平, 崔广禄, 等. SPXY 算法的西瓜可溶性固形物近红外光谱检测[J]. *光谱学与光谱分析*, 2019, 39(3): 738–742. [WANG S F, HAN P, CUI G L, et al. The NIR detection research of soluble solid content in watermelon based on SPXY algorithm[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(3): 738–742.]
- [33] 郭慧娟, 朱思祁, 黎远鹏等. 基于 iPLS 和 SiPLS 算法的人体血清胆红素含量的可见-近红外光谱建模[J]. *光电子激光*, 2016, 27(10): 1136–1144. [GUO H X, ZHU S Q, LI Y P, et al. Visible-near infrared spectroscopy modeling on the contents of serum bilirubin based on iPLS and SiPLS[J]. *Journal of Optoelectronics Laser*, 2016, 27(10): 1136–1144.]
- [34] GUO J, DENG H, CHEN L J, et al. A reliable method for identification of antibiotics by terahertz spectroscopy and svm[J]. *Journal of Spectroscopy*, 2020. <https://doi.org/10.1155/2020/8811467>
- [35] 王欣卉, 宋雪健. 近红外光谱技术对稻花香大米品牌的保护研究[J]. *粮食科技与经济*, 2020, 45(11): 97–99. [WANG X H, SONG X J. Study on the protection of Daohuaxiang rice brand by near infrared spectroscopy[J]. *Grain Science and Technology and Economy*, 2020, 45(11): 97–99.]
- [36] 翁士状, 唐佩佩, 张雪艳, 等. 高光谱成像的图谱特征与卷积神经网络的名优大米无损鉴别[J]. *光谱学与光谱分析*, 2020, 40(9): 2826–2833. [WENG S Z, TANG P P, ZHANG X Y, et al. Non-destructive identification method of famous rice based on image and spectral features of hyperspectral imaging with convolutional neural network[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(9): 2826–2833.]
- [37] 沙敏, 李良翠, 黄家乐, 等. 拉曼光谱数据处理方式对大米产地鉴别模型的影响[J]. *中国食品学报*, 2021, 21(5): 369–376. [SHA M, LI L C, HUANG J L, et al. Effect of data processing method on identification of rice from different geographical origins by raman spectroscopy[J]. *Journal of Chinese Institute of Food Science and Technology*, 2021, 21(5): 369–376.]